Elements of an EmotionML 1.0

W3C Incubator Group Report 20 November 2008

This version:
http://www.w3.org/2005/Incubator/emotion/XGR-emotionml-20081120/
Latest version:
http://www.w3.org/2005/Incubator/emotion/XGR-emotionml/

Editor:
Marc Schröder (DFKI GmbH)
Authors:
(in alphabetic order)
Paolo Baggia (Loquendo, S.p.A.)
Felix Burkhardt (Deutsche Telekom AG)
Jean-Claude Martin (CNRS)
Catherine Pelachaud (CNRS)
Christian Peter (Fraunhofer Gesellschaft)
Björn Schuller (Technische Universität München)
Ian Wilson (Emotion AI)
Enrico Zovato (Loquendo, S.p.A.)

Abstract

As the web is becoming ubiquitous, interactive, and multimodal, technology needs to deal increasingly with human factors, including emotions. The present Final Report of the Emotion Markup Language Incubator Group provides elements for an Emotion Markup Language striking a balance between scientific well-foundedness and practical applicability. The language is conceived as a "plug-in" language suitable for use in three different areas: (1) manual annotation of data; (2) automatic recognition of emotion-related states from user behaviour; and (3) generation of emotion-related system behaviour.

Status of this document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of Final Incubator Group Reports is available. See also the W3C technical reports index at http://www.w3.org/TR/.

The present Final Report consolidates discussions in the Emotion Markup Language Incubator Group (EmotionML XG, 2007-2008) concerning a generally usable markup language for emotions and related states. Earlier work in the Emotion Incubator Group (2006-2007) had identified a comprehensive list of requirements arising from use cases of an Emotion Markup Language. Drawing on this work, the EmotionML XG has identified a set of mandatory requirements, and has started to develop a draft specification for an EmotionML. The present report reflects the degree of consensus that was reached in the Incubator Group. Issue notes are used to highlight open issues and aspects requiring further work.

The present report is conceived as a starting point for future work of a W3C endeavour in the Recommendation Track, in the expectation that it should be possible to develop it into a First Public Working Draft within a very short period of time. The intention of both the EmotionML XG and the MMI WG, agreed at several joint meetings, is to continue the work in the Multimodal Interaction (MMI) Working Group.

This document was developed by the Emotion Markup Language Incubator Group.

Publication of this document by W3C as part of the W3C Incubator Activity indicates no endorsement of its content by W3C, nor that W3C has, is, or will be allocating any resources to the issues addressed by it. Participation in Incubator Groups and publication of Incubator Group Reports at the W3C site are benefits of W3C Membership.

Incubator Groups have as a goal to produce work that can be implemented on a Royalty Free basis, as defined in the W3C Patent Policy. Participants in this Incubator Group have made no statements about whether they will offer licenses according to the licensing requirements of the W3C Patent Policy for portions of this Incubator Group Report that are subsequently incorporated in a W3C Recommendation.

Conventions of this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Table of Contents

1 Introduction

Human emotions are increasingly understood to be a crucial aspect in human-machine interactive systems. Especially for non-expert end users, reactions to complex intelligent systems resemble social interactions, involving feelings such as frustration, impatience, or helplessness if things go wrong. Dealing with these kinds of states in technological systems requires a suitable representation, which should make the concepts and descriptions developed in the scientific literature available for use in technological contexts. To the extent that the web is becoming truly ubiquitous, and involves increasingly multimodal paradigms of interaction, it seems appropriate to define a Web standard for representing emotion-related states, which can provide the required functionality.

This report describes elements of an Emotion Markup Language (EmotionML) designed to be usable in a broad variety of technological contexts while reflecting concepts from the affective sciences.

The report is the result of one year's work in the Emotion Markup Language Incubator Group (EmotionML XG), which built on the results of the Emotion Incubator Group. 21 persons participated in the group: 11 delegates from nine W3C member institutions (Chinese Academy of Sciences, Deutsche Telekom, DFKI, Fraunhofer Gesellschaft, IVML-NTUA, Loquendo, MIMOS BHD, Nuance Communications, and SRI International) as well as ten invited experts. The group worked by consensus where possible; where different options were preferred by different participants, the available choices were identified as such; it was not considered necessary at the stage of an Incubator Group to take final decisions. The specification proposals in this report therefore represent consensus in the group unless noted otherwise; issue notes are used to describe open questions as well as available choices.

1.1 Reasons for defining an Emotion Markup Language

As for any standard format, the first and main goal of an EmotionML is twofold: to allow a technological component to represent and process data, and to enable interoperability between different technological components processing the data.

The Emotion Incubator Group had listed 39 individual use cases for an EmotionML, grouped into three broad types:

1. Manual annotation of material involving emotionality, such as annotation of videos, of speech recordings, of faces, of texts, etc;

2. Automatic recognition of emotions from sensors, including physiological sensors, speech recordings, facial expressions, etc., as well as from multi-modal combinations of sensors;
3. Generation of emotion-related system responses, which may involve reasoning about the emotional implications of events, emotional prosody in synthetic speech, facial expressions and gestures of embodied agents or robots, the choice of music and colors of lighting in a room, etc.

Most of these use cases are still limited to use in research labs, but an increasing number of commercial activities can be observed, both by small startup companies and by larger companies.

Interactive systems are likely to involve both analysis and generation of emotion-related behaviour; furthermore, systems are likely to benefit from data that was manually annotated, be it as training data or for rule-based modelling. Therefore, it is desirable to propose a single EmotionML that can be used in all three contexts.

A second reason for defining an EmotionML is the observation that ad hoc attempts to deal with emotions and related states often lead people to make the same mistakes that others have made before. The most typical mistake is to model emotions as a small number of intense states such as anger, fear, joy, and sadness; this choice is often made irrespective of the question whether these states are the most appropriate for future intended applications. Crucially, the available alternatives that have been developed in the affective science literature are not sufficiently known, resulting in dead-end situations after the initial steps of work. Careful consideration of states to study and of representations for describing them can help avoid such situations.

Given this background, a scientifically-informed EmotionML can help potential users in identifying the suitable representations for their respective applications.

1.2 The challenge of defining a generally usable Emotion Markup Language

Any attempt to standardise the description of emotions using a finite set of fixed descriptors is doomed to failure: even scientists cannot agree on the number of relevant emotions, or on the names that should be given to them. Even more basically, the list of emotion-related states that should be distinguished varies between researchers. Basically, the vocabulary needed depends on the context of use. On the other hand, the basic structure of concepts is less controversial: researchers agree that emotions involve triggers, appraisals, feelings, expressive behaviour including physiological changes, and action tendencies; emotions in their entirety can be described in terms of categories or a small number of dimensions; emotions have an intensity, and so on. For details, see Scientific Descriptions of Emotions in the Final Report of the Emotion Incubator Group.

Given this lack of agreement on descriptors in the field, the only practical way of defining an EmotionML seems to be the definition of possible structural elements, their valid child elements and attributes, but to allow users to "plug in" vocabularies that they consider appropriate for their work. A central repository of such vocabularies can serve as a recommended starting point; where that seems inappropriate, users can create their custom vocabularies.

An additional challenge lies in the aim to provide a generally usable markup, as the requirements arising from the three different use cases (annotation, recognition, and generation) are rather different. Whereas manual annotation tends to require all the fine-grained distinctions considered in the scientific literature, automatic recognition systems can usually distinguish only a very small number of different states. Furthermore, different communities have their deeply engrained customs: for example, when working with scale values, manual annotation generally uses a small number of discrete values on an ordinal scale, whereas machine analysis often produces continuous values.

For the reasons outlined here, it is clear that there is an inevitable tension between flexibility and interoperability, which need to be weighed in the formulation of an EmotionML. The guiding principle in the following specification has been to provide a choice only where it is needed; to propose reasonable default options for every choice; and, ultimately, to propose mapping mechanisms where that is possible and meaningful.

1.3 Glossary of terms

The terms related to emotions are not used consistently, neither in common use nor in the scientific literature. The following glossary attempts to reduce ambiguity by describing the intended meaning of terms in this document.

Action tendency
: Emotions have a strong influence on the motivational state of a subject. Emotion theory associates emotions to a small set of so-called action tendencies, e.g. avoidance (relates to fear), rejecting (disgust) etc. Action tendencies can be viewed as a link between the outcome of an appraisal process and actual actions.

Affect / Affective state
: In the scientific literature, the term "affect" is often used as a general term covering a range of phenomena called "affective states", including emotions, moods, attitudes, etc. Proponents of the term consider it to be more generic than "emotion", in the sense that it covers both acute and long-term, specific and unspecific states. In this report, the term "affect" is avoided so that the scope of the intended markup language is more easily accessible to the non-expert; the term "affective state" is used interchangeably with "emotion-related state".

Appraisal
: The term "appraisal" is used in the scientific literature to describe the evaluation process leading to an emotional response. Triggered by an "emotion-eliciting event", an individual carries out an automatic, subjective assessment of

the event, in order to determine the relevance of the event to the individual. This assessment is carried out along a number of "appraisal dimensions" such as the novelty, pleasantness or goal conduciveness of the event.

Emotion

In this report, the term "emotion" is used in a very broad sense, covering both intense and weak states, short and long term, with and without event focus. This meaning is intended to reflect the understanding of the term "emotion" by the general public. In the scientific literature on emotion theories, the term "emotion" or "fullblown emotion" refers to intense states with a strong focus on current events, often in the context of the survival-benefitting function of behavioural responses such as "fight or flight". This reading of the term seems inappropriate for the vast majority of human-machine interaction contexts, in which more subtle states dominate; therefore, where this reading is intended, the term "fullblown emotion" is used in this report.

Emotion-related state

A cover term for the broad range of phenomena intended to be covered by this specification. In the scientific literature, several kinds of emotion-related or affective states are distinguished, see Emotions and related states in the final report of the Emotion Incubator Group.

Emotion dimensions

A small number of continuous scales describing the most basic properties of an emotion. Often three dimensions are used: valence (sometimes named pleasure), arousal (or activity/activation), and potency (sometimes called control, power or dominance). However, sometimes two, or more than three dimensions are used.

Fullblown emotion

Intense states with a strong focus on current events, often in the context of the survival-benefitting function of behavioural responses such as "fight or flight".

2 Elements of Emotion Markup

The following sections describe the syntax of the main elements of EmotionML as proposed by the EmotionML XG. The specification is not fully complete, but is starting to be sufficiently concrete so that it is possible to see the direction in which the development is going. Feedback is highly appreciated.

2.1 Document structure

2.1.1 Document root: The `<emotionml>` element

Annotation	`<emotionml>`

Definition	The root element of an EmotionML document.
Children	The element MUST contain one or more **<emotion>** elements. It MAY contain a single **<metadata>** element.

Attributes	• **Required**:
	  ○ Namespace declaration for EmotionML, see EmotionML namespace.
	• **Optional**:
	  ○ any other namespace declarations for application-specific namespaces.

Occurrence	This is the root element -- it cannot occur as a child of any other EmotionML elements.

<emotionml> is the root element of a standalone EmotionML document. It wraps a number of **<emotion>** elements into a single document. It may contain a single **<metadata>** element, providing document-level metadata.

The **<emotionml>** element MUST define the EmotionML namespace, and may define any other namespaces.

Example:

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
...
</emotionml>
```

or

```
<em:emotionml xmlns:em="http://www.w3.org/2008/11/emotionml">
...
</em:emotionml>
```

ISSUE: It should be possible to specify, on the document level, default values for the vocabularies used for emotion representations.

Note: One of the envisaged uses of EmotionML is to be used in the context of other markup languages. In such cases, there will be no **<emotionml>** root element, but **<emotion>** elements will be used directly in other markup -- see Examples of possible use with other markup languages.

ISSUE: Should the **<emotionml>** element have a **version** attribute? If so, how would the version of EmotionML used be identified when using **<emotion>** elements directly in other markup?

2.1.2 A single emotion annotation: The <emotion> element

Annotation	<emotion>
Definition	This element represents a single emotion annotation.
Children	All children are optional. If present, the following child elements can occur only once: **<category>**; **<dimensions>**; **<appraisals>**; **<action-tendencies>**; **<intensity>**; **<metadata>**. If present, the following child elements may occur one or more times: **<link>**, **<modality>**. There are no constraints on the combinations of children that are allowed.

Attributes	• **Required**:
	  ◦ None for the moment, but see ISSUE on QNames below.
	• **Optional**:
	  ◦ **`date`**, the absolute date when the annotated event occured.
	  ◦ **`timeRefURI`**, indicating the URI used to anchor a relative timestamp. MUST be given if either **`timeRefAnchor`** or **`offsetToStart`** are specified.
	  ◦ **`timeRefAnchor`**, indicates whether to measure the time from the start or end of the interval designated with **`timeRefURI`**.
	  ◦ **`offsetToStart`**, specifies the offset for the start of input from the anchor point designated with **`timeRefURI`** and **`timeRefAnchor`**.

Occurrence	as a child of **`<emotionml>`**, or in any markup using EmotionML.

The **`<emotion>`** element represents an individual emotion annotation. No matter how simple or complex its substructure is, it represents a single statement about the emotional content of some annotated item. Where several statements about the emotion in a certain context are to be made, several **`<emotion>`** elements MUST be used. See Examples of emotion annotation for illustrations of this issue.

Whereas it is possible to use **`<emotion>`** elements in a standalone **`<emotionml>`** document, a typical use case is expected to be embedding an **`<emotion>`** into some other markup -- see Examples of possible use with other markup languages.

ISSUE: Maybe it should be required that at least one of **`<category>`**, **`<dimensions>`**, **`<appraisals>`** and **`<action-tendencies>`** MUST be present? Otherwise it is possible not to say anything about the emotion as such. Or should **`<intensity>`** be included in this list? Does it make sense to state the intensity of an emotion but not its nature?

ISSUE: If the definition of vocabularies is done using QNames, an optional attribute of the **`<emotion>`** tag may be the namespace definitions for custom vocabularies.

ISSUE: The degree of consensus in the group regarding the name of this element needs clarification. Clarification is also needed regarding the degree of consensus regarding the proposal not to include an attribute 'type' and 'set',

allowing for the explicit annotation of the type of affective state and the indication of the set of possible types of affective state.

2.2 Representations of emotions and related states

2.2.1 The `<category>` element

Annotation	`<category>`
Definition	Description of an emotion or a related state using a single category.
Children	None

Attributes	• **Required**:
	  ◦ `set`, a name or URI identifying the set of category names that can be used.
	  ◦ `name`, the name of the category, which must be contained in the set of categories identified in the `set` attribute.
	• **Optional**:
	  ◦ `confidence`, the annotator's confidence that the annotation is correct.

Occurrence	A single `<category>` MAY occur as a child of `<emotion>`.

`<category>` describes an emotion or a related state in terms of a single category name, given as the value of the `name` attribute. The name MUST belong to a clearly-identified set of category names, which MUST be defined according to Defining vocabularies for representing emotions.

The set of legal values of the `name` attribute is indicated in the `set` attribute of the `<category>` element. Different sets can be used, depending on the requirements of the use case. In particular, different types of emotion-related / affective states can be annotated by using appropriate value sets.

ISSUE: The details of the definition of sets of values need to be sorted out. Throughout this draft, a `set` attribute is used to identify the named set of possible values. Whether a `set` attribute should actually be used, and if so, the

format of its attribute values, needs to be clarified in the context of Defining vocabularies. This issue is related to the section Considerations regarding the validation of EmotionML documents.

Examples:

In the following example, the emotion category "satisfaction" is being annotated; it must be contained in the set of values named "everydayEmotions".

```
<emotion>
    <category set="everydayEmotions" name="satisfaction"/>
</emotion>
```

The following is an annotation of an interpersonal stance "distant" which must belong to the set of values named "commonInterpersonalStances".

```
<emotion>
    <category set="commonInterpersonalStances" name="distant"/>
</emotion>
```

2.2.2 The `<dimensions>` element

Annotation	`<dimensions>`
Definition	Description of an emotion or a related state using a set of dimensions.
Children	`<dimensions>` MUST contain one or more dimension elements. The names of dimension elements which may occur as valid child elements are defined by the `set` attribute.

Attributes	• **Required**:
	  ◦ **`set`**, a name or URI identifying the set of dimension names that can be used.
	• **Optional**:
	  ◦ **`confidence`**, the annotator's confidence that the entirety of dimensional annotation given is correct.

Occurrence	A single **`<dimensions>`** MAY occur as a child of **`<emotion>`**.
Annotation	**Dimension elements**
Definition	Annotation of a single emotion dimension. The tag name must be contained in the list of values identified by the **`set`** attribute of the enclosing **`<dimensions>`** element.
Children	Optionally, a dimension MAY have a **`<trace>`** child element.

Attributes	• **Required**:
	  ◦ (none)
	• **Optional**:
	  ◦ **`value`**, the (constant) scale value of this dimension.
	  ◦ **`confidence`**, the annotator's (constant) confidence that the annotation given for this dimension is correct.

Occurrence	Dimension elements occur as children of **`<dimensions>`**. Valid tag names are constrained to the set of dimension names identified in the **`set`** attribute of the **`<dimensions>`** parent element. For any given dimension name in the set, zero or one occurrences are allowed within a **`<dimensions>`** element.

A **`<dimensions>`** element describes an emotion or a related state in terms of a set of emotion dimensions. The names of the emotion dimensions MUST belong to a clearly-identified set of dimension names, which MUST be defined according to Defining vocabularies for representing emotions.

The set of values that can be used as tag names of child elements of the **`<dimensions>`** element is indicated in the **`set`** attribute of the **`<dimensions>`** element. Different sets can be used, depending on the requirements of the use case.

ISSUE: The details of the definition of sets of values need to be sorted out. Throughout this draft, we assume the use of a **`set`** attribute. Whether a **`set`** attribute should actually be used, and if so, the format of its attribute values, needs to be clarified in the context of Defining vocabularies. This issue is related to the section Considerations regarding the validation of EmotionML documents.

There are no constraints regarding the order of the dimension child elements within a **`<dimensions>`** element.

Any given dimension is either unipolar or bipolar; its **`value`** attribute MUST contain either discrete or continuous Scale values.

ISSUE: the definition of the set of dimensions should include the detailed constraints on valid values of the **`value`** attribute.

A dimension element MUST either contain a **`value`** attribute or a **`<trace>`** child element, corresponding to static and dynamic representations of Scale values, respectively.

If the dimension element has both a **`confidence`** attribute and a **`<trace>`** child, the **`<trace>`** child MUST NOT have a **`samples-confidence`** attribute. In other words, it is possible to either give a constant confidence on the dimension element or a confidence trace on the **`<trace>`** element, but not both.

Examples:

One of the most widespread sets of emotion dimensions used (sometimes by different names) is the combination of valence, arousal and potency. Assuming that arousal and potency are unipolar scales with typical values between 0 and 1, and valence is a bipolar scale with typical values between -1 and 1, the following example is a state of rather low arousal, very positive valence, and high potency -- in other words, a relaxed, positive state with a feeling of being in control of the situation:

```
<emotion>
    <dimensions set="valenceArousalPotency">
        <arousal value="0.3"/><!-- lower-than-average arousal -->
        <valence value="0.9"/><!-- very high positive valence -->
```

```
        <potency value="0.8"/><!-- relatively high potency    -->
    </dimensions>
</emotion>
```

In some use cases, custom sets of application-specific dimensions will be required. The following example uses a custom set of dimensions, defining a single, bipolar dimension "friendliness".

```
<emotion>
    <dimensions set="myFriendlinessDimension">
        <friendliness value="-0.7"/><!-- a pretty unfriendly person -->
    </dimensions>
</emotion>
```

Different use cases require continuous or discrete Scale values; the following example uses discrete values for a bipolar dimension "valence" and a unipolar dimension "arousal".

```
<emotion>
    <dimensions set="discreteValenceArousal">
        <arousal value="very high"/>
        <valence value="slightly negative"/>
    </dimensions>
</emotion>
```

2.2.3 The `<appraisals>` element

Annotation	`<appraisals>`
Definition	Description of an emotion or a related state using appraisal variables.
Children	`<appraisals>` MUST contain one or more appraisal elements. The names of appraisal elements which may occur as valid child elements are identified by the `set` attribute.

Attributes	• **Required**:
	  ◦ **`set`**, a name or URI identifying the set of appraisal names that can be used.
	• **Optional**:
	  ◦ **`confidence`**, the annotator's confidence that the entirety of appraisals annotation given is correct.

Occurrence	A single **`<appraisals>`** MAY occur as a child of **`<emotion>`**.
Annotation	**Appraisal elements**
Definition	Annotation of a single emotion appraisal. The tag name must be contained in the list of values identified by the **`set`** attribute of the enclosing **`<appraisals>`** element.
Children	Optionally, a appraisal MAY have a **`<trace>`** child element.

Attributes	• **Required**:
	  ◦ (none)
	• **Optional**:
	  ◦ **`value`**, the (constant) scale value of this appraisal.
	  ◦ **`confidence`**, the annotator's (constant) confidence that the annotation given for this appraisal is correct.

Occurrence	Appraisal elements occur as children of **`<appraisals>`**. Valid tag names are constrained to the set of appraisal names identified in the **`set`** attribute of the **`<appraisals>`** parent element. For any given appraisal name in the set, zero or one occurrences are allowed within an **`<appraisals>`** element.

An **`<appraisals>`** element describes an emotion or a related state in terms of a set of appraisals. The names of the appraisals MUST belong to a clearly-identified set of appraisal names, which MUST be defined according to Defining vocabularies for representing emotions.

The set of values that can be used as tag names of child elements of the **`<appraisals>`** element is indicated in the **`set`** attribute of the **`<appraisals>`** element. Different sets can be used, depending on the requirements of the use case.

ISSUE: The details of the definition of sets of values need to be sorted out. Throughout this draft, we assume the use of a **set** attribute. Whether a **set** attribute should actually be used, and if so, the format of its attribute values, needs to be clarified in the context of Defining vocabularies. This issue is related to the section Considerations regarding the validation of EmotionML documents.

There are no constraints regarding the order of the appraisal child elements within a **<appraisals>** element.

Any given appraisal is either unipolar or bipolar; its **value** attribute MUST contain either discrete or continuous Scale values.

ISSUE: the definition of the set of appraisals should include the detailed constraints on valid values of the **value** attribute.

An appraisal element MUST either contain a **value** attribute or a **<trace>** child element, corresponding to static and dynamic representations of Scale values, respectively.

If the appraisal element has both a **confidence** attribute and a **<trace>** child, the **<trace>** child MUST NOT have a **samples-confidence** attribute. In other words, it is possible to either give a constant confidence on the appraisal element, or a confidence trace on the **<trace>** element, but not both.

Examples:

One of the most widespread sets of emotion appraisals used is the appraisals set proposed by K. Scherer, namely novelty, intrinsic pleasantness, goal/need significance, coping potential, and norm/self compatibility. Another very widespread set of emotion appraisals, used in particular in computational models of emotion, is the OCC set of appraisals (Ortony et al., 1988), which includes the consequences of events for oneself or for others, the actions of others and the perception of objects. Assuming some appraisal variables, say novelty is a unipolar scale with typical values between 0 and 1, and intrinsic pleasantness is a bipolar scale with typical values between -1 and 1, the following example is a state arising from the evaluation of an unpredicted and quite unpleasant event:

```
<emotion>
```

```
    <appraisals set="Scherer_appraisals_checks">
        <novelty value="0.8"/>
        <intrinsic-pleasantness value="-0.5"/>
    </appraisals>
</emotion>
```

In some use cases, custom sets of application-specific appraisals will be required. The following example uses a custom set of appraisals, defining single, bipolar appraisal "likelihood".

```
<emotion>
    <appraisals set="myLikelihoodAppraisal">
        <likelihood value="0.8"/><!-- a very predictable event -->
    </appraisals>
</emotion>
```

Different use cases require continuous or discrete Scale values; the following example uses discrete values for a bipolar appraisal "intrinsic-pleasantness" and a unipolar appraisal "novelty".

```
<emotion>
    <appraisals set="discreteSchererAppraisals">
        <novelty value="very high"/>
        <intrinsic-pleasantness value="slightly negative"/>
    </appraisals>
</emotion>
```

2.2.4 The `<action-tendencies>` element

Annotation	`<action-tendencies>`
Definition	Description of an emotion or a related state using a set of action tendencies.

Children	**<action-tendencies>** MUST contain one or more action-tendency elements. The names of action-tendency elements which may occur as valid child elements are identified by the **set** attribute.

Attributes	• **Required**:
	  ◦ **set**, a name or URI identifying the set of action-tendency names that can be used.
	• **Optional**:
	  ◦ **confidence**, the annotator's confidence that the entirety of action-tendency annotation given is correct.

Occurrence	A single **<action-tendencies>** MAY occur as a child of **<emotion>**.
Annotation	**Action-tendency elements**
Definition	Annotation of a single action-tendency. The tag name must be contained in the list of values identified by the **set** attribute of the enclosing **<action-tendencies>** element.
Children	Optionally, an action-tendency MAY have a **<trace>** child element.

Attributes	• **Required**:
	  ◦ (none)
	• **Optional**:
	  ◦ **value**, the (constant) scale value of this action-tendency.
	  ◦ **confidence**, the annotator's (constant) confidence that the annotation given for this action-tendency is correct.

Occurrence	action-tendency elements occur as children of **<action-tendencies>**. Valid tag names are constrained to the set of action-tendency names identified in the **set** attribute of the **<action-tendencies>** parent element. For any given action-tendency name in the set, zero or one occurrences are allowed within a **<action-tendencies>** element.

An **<action-tendencies>** element describes an emotion or a related state in terms of a set of action-tendencies. The names of the action-tendencies MUST belong to a clearly-identified set of action-tendency names, which MUST be defined according to Defining vocabularies for representing emotions.

The set of values that can be used as tag names of child elements of the **<action-tendencies>** element is indicated in the **set** attribute of the **<action-tendencies>** element. Different sets can be used, depending on the requirements of the use case.

ISSUE: The details of the definition of sets of values need to be sorted out. Throughout this draft, we assume the use of a **set** attribute. Whether a **set** attribute should actually be used, and if so, the format of its attribute values, needs to be clarified in the context of Defining vocabularies. This issue is related to the section Considerations regarding the validation of EmotionML documents.

There are no constraints regarding the order of the action-tendency child elements within a **<action-tendencies>** element.

Any given action-tendency is either unipolar or bipolar; its **value** attribute MUST contain either discrete or continuous Scale values.

ISSUE: the definition of the set of action-tendencies should include the detailed constraints on valid values of the **value** attribute.

A action-tendency element MUST either contain a **value** attribute or a **<trace>** child element, corresponding to static and dynamic representations of Scale values, respectively.

If the action-tendency element has both a **confidence** attribute and a **<trace>** child, the **<trace>** child MUST NOT have a **samples-confidence** attribute. In other words, it is possible to either give a constant confidence on the action-tendency element, or a confidence trace on the **<trace>** element, but not both.

Examples:

One well known use of action tendencies is by N. Frijda who generally uses the term "action readiness". This model uses a number of action tendencies that are low level, diffuse bahaviors from which more concrete actions could be determined. An example of someone attempting to attract someone they like by being confident, strong and attentive might look like this using unipolar values:

```
<emotion>
    <action-tendencies set="frijdaActionReadiness">
        <approach value="0.7"/><!-- get close -->
        <avoid value="0.0"/>
        <being-with value="0.8"/><!-- be happy -->
        <attending value="0.7"/><!-- pay attention -->
        <rejecting value="0.0"/>
        <non-attending value="0.0"/>
        <agonistic value="0.0"/>
        <interrupting value="0.0"/>
        <dominating value="0.7"/><!-- be assertive -->
        <submitting value="0.0"/>
    </action-tendencies>
</emotion>
```

In some use cases, custom sets of application-specific action-tendencies will be required. The following example shows control values for a robot who works in a factory and uses a custom set of action-tendencies, defining example actions for a robot using bipolar and unipolar values.

```
<emotion>
    <action-tendencies set="myRobotActionTendencies">
        <charge-battery value="0.9"/><!-- need to charge battery soon, be-with
charger -->
        <pickup-boxes value="-0.2"/><!-- feeling tired, avoid work -->
    </action-tendencies>
</emotion>
```

Different use cases require continuous or discrete Scale values; the following example shows control values for a robot who works in a factory and uses discrete values for a bipolar action-tendency "pickup-boxes" and a unipolar action-tendency "seek-shelter".

```
<emotion>
    <action-tendencies set="myRobotActionTendencies">
```

```
        <seek-shelter value="very high"/><!-- started to rain, approach shelter -->
        <pickup-boxes value="slightly negative"/><!-- feeling tired, avoid work -->
    </action-tendencies>
</emotion>
```

2.2.5 The `<intensity>` element

Annotation	`<intensity>`
Definition	Represents the intensity of an emotion.
Children	Optionally, an `<intensity>` element MAY have a `<trace>` child element.

Attributes	• **Required**:
	  ◦ (none)
	• **Optional**:
	  ◦ `value`, the (constant) scale value of the intensity.
	  ◦ `confidence`, the annotator's confidence that the annotation given for this intensity is correct.

Occurrence	One `<intensity>` item MAY occur as a child of `<emotion>`.

`<intensity>` represents the intensity of an emotion. The `<intensity>` element MUST either contain a `value` attribute or a `<trace>` child element, corresponding to static and dynamic representations of scale values, respectively. `<intensity>` is a unipolar scale.

If the `<intensity>` element has both a `confidence` attribute and a `<trace>` child, the `<trace>` child MUST NOT have a `samples-confidence` attribute. In other words, it is possible to either give a constant confidence on the `<intensity>` element, or a confidence trace on the `<trace>` element, but not both.

A typical use of intensity is in combination with `<category>`. However, in some emotion models (e.g. Gebhard, 2005), the emotion's intensity can also be used in combination with a position in emotion dimension space, that is in combination with `<dimensions>`. Therefore, intensity is specified independently of `<category>`.

Example:

A weak surprise could accordingly be annotated as follows.

```
<emotion>
    <intensity value="0.2"/>
    <category set="everydayEmotions" name="surprise"/>
</emotion>
```

The fact that intensity is represented by an element makes it possible to add meta-information. For example, it is possible to express a high **confidence** that the intensity is low, but a low confidence regarding the emotion category, as shown as the last example in the description of **confidence**.

2.3 Meta-information

2.3.1 The confidence attribute

Annotation	confidence
Definition	A representation of the degree of confidence or probability that a certain element of the representation is correct.
Occurrence	An optional attribute of **<category>**, **<dimensions>**, **<appraisals>** and **<action-tendencies>** elements, of dimension, appraisal and action-tendency elements and of **<intensity>**.

Confidence MAY be indicated separately for each of the Representations of emotions and related states. For example, the confidence that the **<category>** is assumed correctly is independent from the confidence that its **<intensity>** is correctly indicated.

Rooted in the tradition of statistics a confidence is usually given in an interval from 0 to 1, resembling a probability. This is an intuitive range opposing e.g. (logarithmic) score values. However, additonally a given yet limited number of discrete values may often be sufficient and more intuitive. Insofar, the confidence is a unipolar Scale value.

Legal values:

- a floating-point value from the interval [0;1];
- a fixed number of discrete values (see ISSUE note of the **value** attribute).

ISSUE: Should legal numeric values be in the range of [0,2] to allow for exaggeration? This would make **confidence** consistent with Scale values.

Examples:

In the following one simple example is provided for each element that MAY carry a **confidence** attribute.

The first example uses a verbal discrete scale value to indicate a very high confidence that surprise is the emotion to annotate.

```
<emotion>
    <category set="everydayEmotions" name="surprise" confidence="++"/>
</emotion>
```

The next example illustrates using continuous scale values for **confidence** to indicate that the annotation of high arousal is probably correct, but the annotation of slightly positive valence may or may not be correct. Note that the choice of verbal vs. numeric scales between the emotion **<dimension>** and its **confidence** is totally independent, i.e. it is fully possible to use verbally specified emotion dimensions with numerically specified **confidence** (as in this example) or any other combination of verbal and numeric scales.

```
<emotion>
    <dimensions set="valenceArousal">
        <arousal value="++" confidence="0.9"/>
        <valence value="+" confidence="0.3"/>
    </dimensions>
</emotion>
```

Accordingly, an example of **<appraisals>** using verbal scales for both the appraisal dimensions themselves and for the confidence. Note that the confidence is always unipolar, but that some of the appraisal dimensions are bipolar.

```
<emotion>
    <appraisals set="Scherer_appraisals_checks">
        <novelty value="++" confidence="+"/>
        <intrinsic-pleasantness value="--" confidence="++"/>
    </appraisals>
</emotion>
```

The example for action tendencies demonstrates an alternative realisation: the example shows confidence as an attribute of the entire group of action tendencies; the confidence indicated (rather high) therefore applies to all action tendencies contained.

```
<emotion>
    <action-tendencies set="approachAvoidFlightFlight" confidence="0.8">
        <approach value="0.9"/>
        <avoid value="0.0"/>
        <flight-flight value="0.9"/>
    </action-tendencies>
</emotion>
```

Finally, an example for the case of **<intensity>**: A high confidence is named that the emotion has a low intensity.

```
<emotion>
    <intensity value="0.1" confidence="0.8"/>
</emotion>
```

Note that, as stated, obviously an emotional annotation can be a combination of some or all of the above, as in the following example: the intensity of the emotion is quite probably low, but if we have to guess, we would say the emotion is boredom.

```
<emotion>
    <intensity value="0.1" confidence="0.8"/>
    <category set="everydayEmotions" name="boredom" confidence="0.1"/>
</emotion>
```

ISSUE: It remains to be decided whether **confidence** shall be allowed as attribute of global metadata. Similarily, it has to remain open whether it may be an attribute to complex emotions and regulation, being open by themselves at present.

Further, a tag might be needed to link a **confidence** to a method by which it has been determined, given that emotion recognition systems may use several methods for determining confidence in parallel.

2.3.2 The <modality> element

Annotation	<modality>
Definition	Element used for the annotation of modality.
Children	None

Attributes	• **Required**:
	  ○ **set**, a name or URI identifying the set of modality names that can be used.
	  ○ **mode**, the name of the modality, which must be contained in the set of modalities identified in the **set** attribute.
	• **Optional**:
	  ○ **medium**, the name of the medium through which the emotion has been observed. It must be contained in the set of media named in the **set** attribute.

Occurrence	This element MAY occur as a child of any **<emotion>** element.

The **<modality>** element is used to annotate the modes in which the emotion is reflected. The **mode** attribute can contain

values from a closed set of values, namely those specified by the **set** attribute. For example, a basic or default set could include values like face, voice, body and text. The **mode** and **medium** attributes can contain a list of space separated values, in order to indicate multimodal input or output.

ISSUE: Standard values should be defined for the **mode** and **medium** attributes. For **mode**, common values are "voice", "face", "body", and "text". For **medium** those from EMMA could be used: "acoustic", "visual" and "tactile", complemented by "infrared" for infrared cameras and "bio" or "physio" for physiological readings (to be discussed).

The advantages of including a medium attribute, at the cost of a more complex syntax, are:

- The possibility of annotating modalities which are observed through different media
- The possibility of annotating modalities with two different levels of detail
- The use of this attribute to group more modalities into broader classes for processing reasons.

Example:

In the following example the emotion is expressed through the voice, which is a modality included in the basicModalities set.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
    <emotion>
        <category set="everydayEmotions" name="satisfaction"/>
        <modality set="basicModalities" mode="voice"/>
    </emotion>
</emotionml>
```

In case of multimodal expression of an emotion, a list of space separated modalities can be indicated in the mode attribute, like in the following example in which the two values "face" and "voice" must be included in the basicModalities set.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
    <emotion>
        <category set="everydayEmotions" name="satisfaction"/>
        <modality set="basicModalities" mode="face voice"/>
```

```
    </emotion>
</emotionml>
```

See also the example at section 5.1.2 Automatic recognition of emotions

ISSUE: An alternative way of representing more modalities is to indicate one **<modality>** element for each of them. In order to better classify and distinguish them, an identifier attribute could be introduced.

ISSUE: Depending on the previous issue, it must be discussed whether one or more than one **<modality>** elements can occur inside an **<emotion>** element.

2.3.3 The <metadata> element

Annotation	<metadata>
Definition	This element can be used to annotate arbitrary metadata.
Occurence	A single **<metadata>** elements MAY occur as a child of the **<emotionml>** root tag to indicate global metadata, i.e. the annotations are valid for the document scope; furthermore, a single **<metadata>** element MAY occur as a child of each **<emotion>** element to indicate local metadata that is only valid for that **<emotion>** element.

This element can contain arbitrary data (one option could be [RDF] data), either on a document global level or on a local "per annotation element" level.

ISSUE: One of the design goals of the EmotionML was that no free text should be allowed, i.e. if all XML tags would be removed from a document, no annotation would be left. This current **<metadata>** element enables a violation of this rule.
An alternative to annotate metadata would be a generic element that uses "name" and "value" attributes to express arbitrary data, though this would be less flexible, because complex data structures (like XML data) could not be used.

Examples:

In the following example, the automatic classification for an annotation document was performed by a classifier based on Gaussian Mixture Models (GMM); the speakers of the annotated elements were of different German origins.

```
<emotionml>
    <metadata>
        <classifiers:classifier classifiers:name="GMM"/>
    </metadata>
    <emotion>
        <metadata>
            <origin:localization value="bavarian"/>
        </metadata>
        <category set="everydayEmotions" name="joy"/>
    </emotion>
    <emotion>
        <metadata>
            <origin:localization value="swabian"/>
        </metadata>
        <category set="everydayEmotions" name="sadness"/>
    </emotion>
</emotionml>
```

2.4 Links and time

2.4.1 The `<link>` element

Annotation	`<link>`
Definition	Links may be used to relate the emotion annotation to the "rest of the world", more specifically to the emotional expression, the experiencing subject, the trigger, and the target of the emotion.
Children	None

Attributes	• **Required**:
	  ◦ **`uri`**, a URI identifying the actual link.
	• **Optional**:
	  ◦ **`role`**, the type of relation between the emotion and the external item referred to; one of "expressedBy" (default), "experiencedBy", "triggeredBy", "targetedAt".
	  ◦ **`start`** denotes the start timepoint of an emotion display in a media file. It defaults to "0".
	  ◦ **`end`** denotes the end timepoint of an emotion display in a media file. It defaults to the time length of the media file.

Occurrence	Multiple **`<link>`** items MAY occur as children of **`<emotion>`**.

A **`<link>`** element provides a link to media as a URI [RFC3986]. The semantics of links are described by the **`role`** attribute which MUST have one of four values:

- "expressedBy" indicates that the link points to observable behaviour expressing the emotion. This is the default value if the **`role`** attribute is not explicitly stated;
- "experiencedBy" indicates that the link refers to the subject experiencing the emotion;
- "triggeredBy" indicates that the link identifies an emotion-eliciting event that caused the emotional reaction;
- "targetetAt" indicates that the link points to an object towards which an emotion-related action, or action tendency, is directed.

For resources representing a period of time, start and end time MAY be denoted by use of the optional attributes **`start`** and **`end`** that default to "0" and the time length of the media file, respectively.

ISSUE: What do the default values of "start" and "end" mean for resources that do not have a notion of time, such as XML nodes, picture files, etc.? Maybe there should not be default values, so start and end are unspecified if the **`start`** and **`end`** attributes are not explicitly stated?

There is no restriction regarding the number of **`<link>`** elements that MAY occur as children of **`<emotion>`**.

Example:

The following example illustrates the link to two different URIs having a different **role** with respect to the emotion: one link points to the emotion's expression, e.g. a video clip showing a user expressing the emotion; the other link points to the trigger that caused the emotion, e.g. another video clip that was seen by the person eliciting the expressed emotion. Note that no media sub-classing is used to differentiate between different media types as audio, video, text, etc. Several links may follow as children of one **<emotion>** tag, even having the same **role**: for example a video and physiological sensor data of the expressed emotion.

```
<emotion>
    <link uri="http:..." role="expressedBy"/>
    <link uri="http:..." role="triggeredBy"/>
</emotion>
```

ISSUE: Position on a time line in externally linked objects needs to be finalised.

Agreement was found to include absolute and relative timing. Start and end provision is preferred over provision of a duration attribute. Further no onset, hold, or decay will be included at the moment. However, the following questions remain:

- How should timing be defined syntactically?
- It needs to be specified where timing may occur, that is, is it an element or an attribute (as presently contained by **start** and **end**). Thereby only one of these choices should exist.

2.4.2 Timestamps

2.4.2.1 Absolute time

Annotation	**date**
Definition	Attribute to denote an absolute timepoint as specified in the ISO-8601 standard.
Occurence	The attribute MAY occur inside an **<emotion>** element.

date denotes the absolute timepoint at which an emotion or related state happened. This might be used for example with an "emotional diary" application. The attribute MAY be used with an **<emotion>** element, and MUST be a string in conformance to W3C datetime note based on the ISO-8601 standard.

ISSUE: How to specify dates before christ?

Examples:

In the following example, the emotion category "joy" is annotated for the 23 November 2001, 14:36 hours UTC.

```
<emotion date="2001-11-23T14:36Z">
    <category set="everydayEmotions" name="joy"/>
</emotion>
```

2.4.2.2 Timing in media

Annotation	start, end
Definition	Attributes to denote start and endpoint of an annotation in a media stream. Allowed values must be conform with the SMIL clock value syntax
Occurence	The attributes MAY occur inside a **<link>** element.

start denotes the timepoint from which on an emotion or related state is displayed in a media file. It is optional and defaults to "0".

end denotes the timepoint at which an emotion or related state ends to be displayed in a media file. It is optional and defaults to the time length of the media file.

Both attributes MAY be used with a **<link>** element and MUST be a string in conformance to the SMIL clock value syntax.

ISSUE: Is the SMIL clock value syntax too complicated and should be replaced by simple milliseconds as used in EMMA?

Examples:

In the following example, the emotion category "joy" is diplayed in a video file called "myVideo.avi" from the 3rd to the 9th second.

```
<emotion>
    <category set="everydayEmotions" name="joy"/>
    <link uri="myVideo.avi" start="3s" end="9s"/>
</emotion>
```

2.4.2.3 Timing reference

Annotation	**timeRefURI**
Definition	Attribute indicating the URI used to anchor the relative timestamp.
Annotation	**timeRefAnchor**
Definition	Attribute indicating whether to measure the time from the start or end of the interval designated with **timeRefURI**. Possible values are "start" and "end", default value is "start".
Annotation	**offsetToStart**
Definition	Attribute with a time value, defaulting to zero. It specifies the offset for the start of input from the anchor point designated with **timeRefURI** and **timeRefAnchor**. Allowed values must be conform with the SMIL clock value syntax
Occurence	The above attributes MAY occur as part of an **<emotion>**. If **offsetToStart** or **timeRefAnchor** are given, **timeRefURI** MUST also be specified.

timeRefURI, **timeRefAnchor** and **offsetToStart** may be used to set the timing of an emotion or related state relative to the timing of another annotated element.

ISSUE: Is the SMIL clock value syntax too complicated and should be replaced by simple milliseconds as used in EMMA?

Examples:

In the following example, Fred is annotated as being sad on 23 November 2001 at 14:39 hours, three minutes later than the absolutely positioned reference element.

```
<emotion id="annasJoy" date="2001-11-23T14:36Z">
    <category set="everydayEmotions" name="joy"/>
</emotion>
<emotion id="fredsSadness" timeRefURI="#annasJoy"
    timeRefAnchor="end" offsetToStart="3min">
    <category set="everydayEmotions" name="sadness"/>
</emotion>
```

ISSUE: Is it important to provide for explicit elements or attributes to annotate onset, hold and decay phases? Here's an example for a possible syntax:

<emoml:timing>
<emoml:onset start="00:00:01:00" duration="00:00:04:00" />
<emoml:hold start="00:00:05:00" duration="00:00:02:00" />
<emoml:decay start="00:00:07:00" duration="00:00:06:00" />
</emoml:timing>

2.5 Scale values

Scale values are needed to represent content in dimension, appraisal and action-tendency elements, as well as in

<intensity> and **confidence**.

Representations of scale values can vary along three axes:

- static vs. dynamic: a static, constant scale value is represented using the **value** attribute; for dynamic values, their evolution over time is expressed using the **<trace>** element.
- unipolar vs. bipolar: conceptually, a scale can represent concepts that vary from "nothing" to "a lot" (unipolar scales), or concepts that vary between two opposites, from "very negative" to "very positive" (bipolar scales).
- numeric vs. discrete: some use cases require scale values to be represented as continuous numeric values, whereas other use cases require a number of discrete values.

2.5.1 The value attribute

Annotation	value
Definition	Representation of a static scale value.
Occurrence	An optional attribute of dimension, appraisal and action-tendency elements and of **<intensity>**; these elements MUST either contain a **value** attribute or a **<trace>** element.

The **value** attribute represents a static scale value of the enclosing element.

Conceptually, each dimension, appraisal and action-tendency element is either unipolar or bipolar. The definition of a set of dimensions, appraisals or action tendencies MUST define, for each item in the set, whether it is unipolar or bipolar.

<intensity> is a unipolar scale.

Legal values:

- For unipolar scales, legal values are one of
  - a floating-point value from the interval [0;2], where usual values are in the range [0;1], and values in [1;2] can be used to represent exaggerated values;
  - a fixed number of discrete values (see ISSUE note below).
- For bipolar scales, legal values are one of

- a floating-point value from the interval [-2;2], where usual values are in the range [-1;1], and values in [-2;-1] and [1;2] can be used to represent exaggerated values;
- a fixed number of discrete values (see ISSUE note below).

ISSUE: The list of legal discrete values needs to be finalised. There are two options for discrete five-point scales:

- verbal scales, such as “very negative – negative – neutral – positive – very positive”;
- abstract scales, such as “-- - 0 + ++”

It seems difficult to find generic wordings for verbal scales which fit to all possible uses; however, abstract scales may be unintuitive to use. One option would be to use the definition of vocabulary sets for dimensions, appraisals and action tendencies to define the list of legal discrete values for each dimension. As a result, there would potentially be different discrete values, potentially even a different number of values, for each dimension. Generic interpretability may still be possible, though, because of the requirement to state whether a scale is unipolar or bipolar and in combination with a requirement to list the possible values in increasing order.

ISSUE: Should we allow users to define a different range of legal numeric values, e.g. [0;8], related to a Likert scale?

Examples of the **`value`** attribute can be found in the context of the dimension, appraisal and action-tendency elements and of **`<intensity>`**.

2.5.2 The `<trace>` element

Annotation	`<trace>`
Definition	Representation of the time evolution of a dynamic scale value.
Children	None

Attributes	• **Required**: ◦ **`freq`**, a sampling frequency in Hz. ◦ **`samples`**, a space-separated list of numeric scale values representing the scale value of the enclosing element as it changes over time. • **Optional**: ◦ **`samples-confidence`**, a space-separated list of numeric scale values representing the annotator's confidence that the annotation is correct, as it changes over time.
Occurrence	An optional child element of dimension, appraisal and action-tendency elements and of **`<intensity>`**; these elements MUST either contain a **`value`** attribute or a **`<trace>`** element.

A **`<trace>`** element represents the time course of a numeric scale value. It cannot be used for discrete scale values.

The **`freq`** attribute indicates the sampling frequency at which the values listed in the **`samples`** attribute are given.

A **`<trace>`** MAY include a trace of the confidence alongside with the trace of the scale itself, in the **`samples-confidence`** attribute. If present, **`samples-confidence`** MUST use the same sampling frequency as the content scale, as given in the **`freq`** attribute. If the enclosing element contains a (static) **`confidence`** attribute, the **`<trace>`** MUST NOT have a **`samples-confidence`** attribute. In other words, it is possible to indicate either a static or a dynamic confidence for a given scale value, but not both.

NOTE: The **`<trace>`** representation requires a periodic sampling of values. In order to represent values that are sampled aperiodically, separate **`<emotion>`** annotations with appropriate timing information and individual **`value`** attributes may be used.

Examples:

The following example illustrates the use of a trace to represent an episode of fear during which intensity is rising, first gradually, then quickly to a very high value. Values are taken at a sampling frequency of 10 Hz, i.e. one value every 100 ms.

```
<emotion>
```

```
    <category set="everydayEmotions" name="fear"/>
    <intensity>
        <trace freq="10Hz" samples="0.1 0.1 0.15 0.2 0.2 0.25 0.25 0.25 0.3 0.3 0.35
0.5 0.7 0.8 0.85 0.85"/>
    </intensity>
</emotion>
```

The following example combines a trace of the appraisal "novelty" with a global confidence that the values represent the facts properly. There is a sudden peak of novelty; the annotator is reasonable certain that the annotation is correct:

```
<emotion>
    <appraisals set="someSetWithNovelty">
        <novelty confidence="0.75">
            <trace freq="10Hz" samples="0.1 0.1 0.1 0.1 0.1 0.7 0.8 0.8 0.8 0.8 0.4
0.2 0.1 0.1 0.1"/>
        </novelty>
    </appraisals>
</emotion>
```

In the following example, the confidence itself also changes over time. The observation is the same as before, but the confidence drops at the point where the novelty is rising, indicating some uncertainty where exactly the novelty appraisal is rising:

```
<emotion>
    <appraisals set="someSetWithNovelty">
        <novelty>
            <trace freq="10Hz" samples="0.1 0.1 0.1 0.1 0.1 0.7 0.8 0.8 0.8 0.8 0.4
0.2 0.1 0.1 0.1"
                   samples-confidence="0.7 0.7 0.7 0.4 0.3 0.3 0.3 0.7 0.7 0.7 0.7
0.7 0.7 0.7 0.7"/>
        </novelty>
    </appraisals>
</emotion>
```

3 Defining vocabularies for representing emotions

EmotionML markup MUST refer to one or more vocabularies to be used for representing emotion-related states. Due to the lack of agreement in the community, the EmotionML specification does not preview a single default set which should apply if no set is indicated. Instead, the user MUST explicitly state the value set used.

ISSUE: How to define the actual vocabularies to use for **<category>**, **<dimensions>**, **<appraisals>** and **<action-tendencies>** remains to be specified. As described in Considerations regarding the validation of EmotionML documents, a suitable method may be to define an XML format in which these sets can be defined. The format for defining a vocabulary MUST fulfill at least the following requirements:

- it MUST be possible to refer to a vocabulary, e.g. by means of a URI;
- the vocabulary MUST state explicitly whether it represents category names, dimension elements, appraisal elements or action tendency elements;
- the use of the vocabulary MUST only be possible for the intended type of use (e.g., it MUST NOT be possible to use category names as dimension elements);
- for dimension, appraisal and action tendency vocabularies, the legal Scale values of each dimension MUST be defined.

Furthermore, the format SHOULD allow for

- formal or informal annotation of the meaning of the vocabulary set as a whole (e.g., the type of affective state being described) and
- formal or informal annotation of the meaning of the vocabulary items used (characterisation of the meaning of a vocabulary item, either by informal description or by using some formal ontology).

3.1 Centrally defined default vocabularies

ISSUE: The EmotionML specification SHOULD come with a carefully-chosen selection of default vocabularies,

representing a suitably broad range of emotion-related states and use cases. Advice from the affective sciences SHOULD be sought to obtain a balanced set of default vocabularies.

3.2 User-defined custom vocabularies

EmotionML markup makes no syntactic difference between referring to centrally-defined default vocabularies and referring to user-defined custom vocabularies. Therefore, one option to define a custom vocabulary is to create a definition XML file in the same way as it is done for the default vocabularies.

ISSUE: In addition, it may be desirable to embed the definition of custom vocabularies inside an **<emotionml>** document, e.g. by placing the definition XML element as a child element below the document element **<emotionml>**.

4 Conformance

4.1 EmotionML namespace

The EmotionML namespace is "http://www.w3.org/2008/11/emotionml". All EmotionML elements MUST use this namespace.

ISSUE: This section is a stub. It will be filled with the proper content in a future working draft.

4.2 Use with other namespaces

The EmotionML namespace is intended to be used with other XML namespaces as per the Namespaces in XML Recommendation (1.0 [XML-NS10] or 1.1 [XML-NS11], depending on the version of XML being used).

ISSUE: This section is a stub. It will be filled with the proper content in a future working draft.

4.3 Considerations regarding the validation of EmotionML documents

There is an intrinsic tension between the requirement of using plug-in vocabularies and the formal verification that a document is valid with respect to the specification. The issue has been pointed out repeatedly throughout this report, and is not yet solved. The following two subsections provide elements which may be part of a solution.

4.3.1 Use of QNAMES

A proposal under consideration is to use QNAMES to specify custom values for attributes. This solution allows to substitute the **set** attribute from many elements with a namespace declaration to be used as QNAME for the value of the attribute.

With this solution the attribute values are one or more white space separated QNames as defined in Section 4 of Namespaces in XML (1.0 [XML-NS10] or 1.1 [XML-NS11], depending on the version of XML being used).

When the attribute content is a QName, it is expanded into an expanded-name using the namespace declarations that are in scope for the relative element. Thus, each QName provides a reference to a specific item in the referred namespace.

In the example below, the QName "everydayEmotions:satisfaction" is the value of the **name** attribute and it will be expanded to the "satisfaction" item in the "http://www.example.com/everyday_emotion_catg_tags" namespace. The taxonomy for the everyday emotion categories has to be documented at the specified namespace URI.

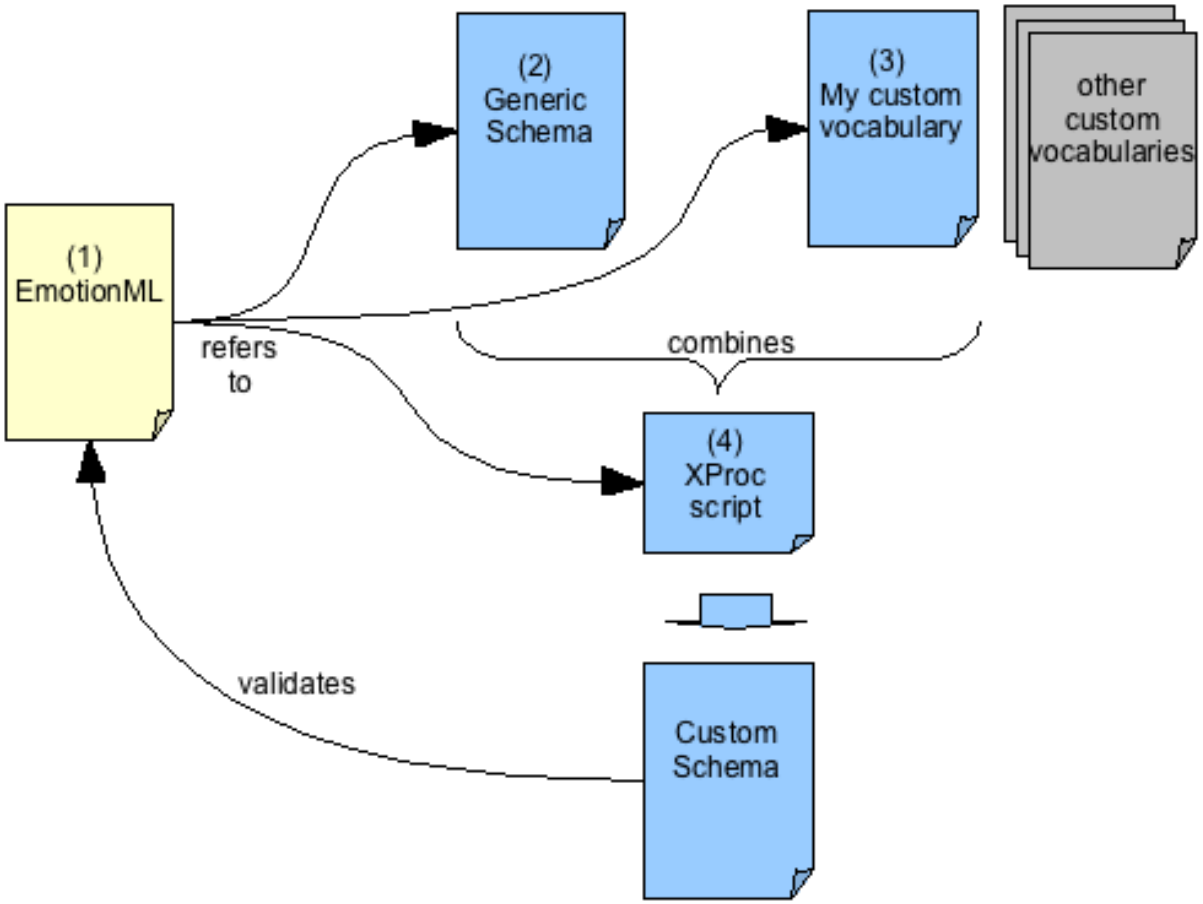
```
<emotionml
  xmlns="http://www.w3.org/2008/11/emotionml"
  xmlns:everydayEmotions="http://www.example.com/everyday_emotion_catg_tags">
    <emotion>
        <category name="everydayEmotions:satisfaction"/>
    </emotion>
</emotionml>
```

This solution allows for referencing different dictionaries depending on the namespace declarations. Moreover, the namespace qualification will make the new set of values unique. The drawbacks of this solution are the absence of a simple and clear way on how to validate the QNAME attribute values, and a more verbose syntax of the attribute contents.

4.3.2 Dynamic schema creation

A static schema document can only fully validate a language where the vaild element names and attribute values are known at the time when the schema is written. For EmotionML, this is not possible because of the fundamental requirement to give users the option of using their own vocabularies.

The following is an idea for dynamically creating a schema document from a base schema and the vocabulary sets referenced in the document itself.



1. An EmotionML document refers to a centrally defined generic schema, one or more vocabularies which may be centrally defined or user-specific, and to a centrally defined XProc script;

2. the generic schema defines the legal structure of EmotionML documents using a schema language such as XML Schema, RelaxNG or Schematron, but using placeholders for concrete vocabulary items;
3. each vocabulary is defined using an XML format known to the XProc script;
4. the XProc script defines the workflow for validating an EmotionML document:
    1. from the EmotionML document, look up the generic schema and the custom vocabularies;
    2. through a suitable mechanism such as XSLT, merge the generic schema and the custom vocabularies into a (short-lived) custom schema;
    3. validate the EmotionML document using the custom schema in the usual way.

ISSUE: The choice of a suitable schema language depends on the required expressive power. The schema language must allow for the verification of both attribute values (for **`<category>`**) and child element names (for **`<dimensions>`**, **`<appraisals>`** and **`<action-tendencies>`**) in a given set, which is either identified using the **`set`** attribute or using QNAMES.

ISSUE: It is unclear how user software can know that an EmotionML document is to be validated using the XProc script.

5 Examples

5.1 Examples of emotion annotation

5.1.1 Manual annotation of emotional material

Use case 1b-ii : Annotation of static images

An image gets annotated with several emotion categories at the same time, but different intensities.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
   <metadata>
      <media-type>image</media-type>
      <media-id>disgust</media-id>
      <media-set>JACFEE-database</media-set>
```

```
      <doc>Example adapted from (Hall & Matsumoto 2004) http://www.davidmatsumoto.
info/Articles/2004_hall_and_matsumoto.pdf
      </doc>
   </metadata>
   <emotion>
       <category set="basicEmotions" name="Disgust"/>
       <intensity value="0.82"/>
   </emotion>
   <emotion>
       <category set="basicEmotions" name="Contempt"/>
       <intensity value="0.35"/>
   </emotion>
   <emotion>
       <category set="basicEmotions" name="Anger"/>
       <intensity value="0.12"/>
   </emotion>
   <emotion>
       <category set="basicEmotions" name="Surprise"/>
       <intensity value="0.53"/>
   </emotion>
</emotionml>
```

Use case 1c-i : Annotation of videos

Example 1: Annotation of a whole video: several emotions are annotated with different intensities.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
    <metadata>
        <media-type>video</media-type>
        <media-name>ed1_4</media-name>
        <media-set>humaine database</media-set>
        <coder-set>JM-AB-UH</coder-set>
    </metadata>
    <emotion>
```

```
        <category set="humaineDatabaseLabels" name="Amusement"/>
        <intensity value="0.52"/>
    </emotion>
    <emotion>
        <category set="humaineDatabaseLabels" name="Irritation"/>
        <intensity value="0.63"/>
    </emotion>
    <emotion>
        <category set="humaineDatabaseLabels" name="Relaxed"/>
        <intensity value="0.02"/>
    </emotion>
    <emotion>
        <category set="humaineDatabaseLabels" name="Frustration"/>
        <intensity value="0.87"/>
    </emotion>
    <emotion>
        <category set="humaineDatabaseLabels" name="Calm"/>
        <intensity value="0.21"/>
    </emotion>
    <emotion>
        <category set="humaineDatabaseLabels" name="Friendliness"/>
        <intensity value="0.28"/>
    </emotion>
</emotionml>
```

Example 2: Annotation of a video segment, where two emotions are annotated for the same timespan.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
    <metadata>
        <media-type>video</media-type>
        <media-name>ext-03</media-name>
        <media-set>EmoTV</media-set>
        <coder>4</coder>
    </metadata>
    <emotion>
```

```
        <category set="emoTV-labels" name="irritation"/>
        <intensity value="0.46"/>
        <link uri="ext03.avi" start="3.24s" end="15.4s">
    </emotion>
    <emotion>
        <category set="emoTV-labels" name="despair"/>
        <intensity value="0.48"/>
        <link uri="ext03.avi" start="3.24s" end="15.4s"/>
    </emotion>
</emotionml>
```

5.1.2 Automatic recognition of emotions

This example shows how automatically annotated data from three affective sensor devices might be stored or communicated.

It shows an excerpt of an episode experienced on 23 November 2001 from 14:36 onwards. Each device detects an emotion, but at slightly different times and for different durations.

The next entry of observed emotions occurs about 6 minutes later. Only the physiology sensor has detected a short glimpse of anger, for the visual and IR camera it was below their individual threshold so no entry from them.

For simplicity, all devices use categorical annotations and the same set of categories. Obviously it would be possible, and even likely, that different devices from different manufacturers provide their data annotated with different emotion sets.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
...
<emotion date="2001-11-23T14:36Z">
  <!--the first modality detects excitement.
      It is a camera observing the face. An URI to the database
      (a dedicated port at the server) is provided to access the
      video stream.-->
  <category set="everyday" name="excited"/>
  <modality medium="visual" mode="face"/>
```

```
  <link uri="http://192.168.1.101:456" start="26s" end="98s"/>
</emotion>

<emotion date="2001-11-23T14:36Z">
  <!--the second modality detects anger. It is an IR camera
      observing the face. An URI to the database (a dedicated port
      at the server) is provided to access the video stream.-->
  <category set="everyday" name="angry"/>
  <modality medium="infrared" mode="face"/>
  <link uri="http://192.168.1.101:457" start="23s" end="108s"/>
</emotion>

<emotion date="2001-11-23T14:36Z">
  <!--the third modality detects excitement again. It is a
      wearable device monitoring physiological changes in the
      body. An URI to the database (a dedicated port at the
      server) is provided to access the data stream.-->
  <category set="everyday" name="excited"/>
  <modality medium="physiological" mode="body"/>
  <link uri="http://192.168.1.101:458 start="19s" end="101s"/>
</emotion>

<emotion date="2001-11-23T14:42Z">
  <category set="everyday" name="angry"/>
  <modality medium="physiological" mode="body"/>
  <link uri="http://192.168.1.101:458 start="2s" end="6s"/>
</emotion>
...
</emotionml>
```

NOTE that handling of complex emotions is not yet specified, see Complex emotions. This example assumes that parallel occurences of emotions will be determined on the time stamp.

NOTE that the used set of emotion descriptions needs to be specified for the document, see Defining vocabularies for representing emotions.

ISSUE: This example assumes that time information at the emotion element level is in full minutes and the modality specific times ofsets on this. This needs to be specified and consistent throughout the markup.

5.1.3 Generation of emotion-related system behaviour

The following example describes various aspects of an emotionally competent robot.

```
<emotionml xmlns="http://www.w3.org/2008/11/emotionml">
    <metadata>
        <name>robbie the robot example</name>
    </metadata>

    <!-- Appraised value of incoming event -->
    <emotion>
        <modality mode="senses"/>
        <appraisals set="scherer_appraisals_checks">
            <novelty value="0.8" confidence="0.4"/>
            <intrinsic-pleasantness value="-0.5" confidence="0.8"/>
        </appraisals>
    </emotion>

    <!-- Robots current internal state configuration -->
    <emotion>
        <modality mode="internal"/>
        <dimensions set="arousal_valence_potency">
            <arousal value="0.3"/>
            <valence value="0.9"/>
            <potency value="0.8"/>
        </dimensions>
    </emotion>

    <!-- Robots output action tendencies -->
    <emotion>
```

```
        <modality mode="body"/>
        <action-tendencies set="myRobotActionTendencies">
            <charge-battery value="0.9"/>
            <seek-shelter value="0.7"/>
            <pickup-boxes value="-0.2"/>
        </action-tendencies>
    </emotion>

    <!-- Robots facial gestures -->
    <emotion>
        <modality mode="face"/>
        <category set="ekman_universal" name="joy"/>
        <link role="expressedBy" start="0" end="5s" uri="smile.xml"/>
    </emotion>
</emotionml>
```

5.2 Examples of possible use with other markup languages

One intended use of EmotionML is as a plug-in for existing markup languages. For compatibility with text-annotating markup languages such as SSML, EmotionML avoids the use of text nodes. All EmotionML information is encoded in element and attribute structures.

This section illustrates the concept using two existing W3C markup languages: EMMA and SSML.

5.2.1 Use with EMMA

EMMA is made for representing arbitrary analysis results; one of them could be the emotional state. The following example represents an analysis of a non-verbal vocalisation; its emotion is described as most probably a low-intensity state, maybe boredom.

```
<emma:emma version="1.0" xmlns:emma="http://www.w3.org/2003/04/emma"
        xmlns="http://www.w3.org/2008/11/emotionml">
    <emma:interpretation start="12457990" end="12457995" mode="voice" verbal="false">
```

```
        <emotion>
            <intensity value="0.1" confidence="0.8"/>
            <category set="everydayEmotions" name="boredom" confidence="0.1"/>
        </emotion>

    </emma:interpretation>
</emma:emma>
```

5.2.2 Use with SSML

Two options for using EmotionML with SSML can be illustrated.

First, it is possible with the current draft version of SSML [SSML 1.1] to use arbitrary markup belonging to a different namespace anywhere in an SSML document; only SSML processors that support the markup would take it into account. Therefore, it is possible to insert EmotionML below, for example, an **<s>** element representing a sentence; the intended meaning is that the enclosing sentence should be spoken with the given emotion, in this case a moderately doubtful tone of voice:

```
<?xml version="1.0"?>
<speak version="1.1" xmlns="http://www.w3.org/2001/10/synthesis"
         xmlns:emo="http://www.w3.org/2008/11/emotionml"
         xml:lang="en-US">
    <s>
        <emo:emotion>
            <emo:category set="everydayEmotions" name="doubt"/>
            <emo:intensity value="0.4"/>
        </emo:emotion>

        Do you need help?
    </s>
</speak>
```

Second, a future version of SSML could explicitly preview the annotation of paralinguistic information, which could fill the

gap between the extralinguistic, speaker-constant settings of the **<voice>** tag and the linguistic elements such as **<s>**, **<emphasis>**, **<say-as>** etc. The following example assumes that there is a **<style>** tag for paralinguistic information in a future version of SSML. The style could either embed an **<emotion>**, as follows:

```
<?xml version="1.0"?>
<speak version="x.y" xmlns="http://www.w3.org/2001/10/synthesis"
         xmlns:emo="http://www.w3.org/2008/11/emotionml"
         xml:lang="en-US">
    <s>
      <style>
        <emo:emotion>
            <emo:category set="everydayEmotion" name="doubt"/>
            <emo:intensity value="0.4"/>
        </emo:emotion>

        Do you need help?
      </style>
    </s>
</speak>
```

Alternatively, the **<style>** could refer to a previously defined **<emotion>**, for example:

```
<?xml version="1.0"?>
<speak version="x.y" xmlns="http://www.w3.org/2001/10/synthesis"
         xmlns:emo="http://www.w3.org/2008/11/emotionml"
         xml:lang="en-US">
    <emo:emotion id="somewhatDoubtful">
        <emo:category set="everydayEmotion" name="doubt"/>
        <emo:intensity value="0.4"/>
    </emo:emotion>

    <s>
        <style ref="#somewhatDoubtful">
            Do you need help?
```

```
        </style>
    </s>
</speak>
```

6 Comments and general open issues

6.1 Complex emotions

Emotions can be complex, in the sense that multiple emotions are present in a single emotional episode. Individual emotions may occur in different intensities, resulting in a blend or mixture of emotions; for example, it is possible to be slightly surprised yet happy, or slightly depressed and strongly angry. Co-occurrence of several emotions may be due to different causes of an emotion and/or differences in expression in different modalities. With the current specification of EmotionML, such cases can be represented using several **<emotion>** elements with different **<link role="triggeredBy".../>** or different **<modality>** elements. It is not yet possible to represent complex emotions involving regulation.

Example:

The following examples illustrates how it is possible with the present specification to describe a complex emotion, consisting of two different categories.

In the first example, the two emotions are simultaneously expressed, but triggered by different causes. The two emotions are experienced by "John", and expressed at the same time. The emotion "satisfaction" is triggered by the number of guests on the guest list; the emotion "surprise" is triggered by an event recorded in the same video file as John's expression (the fact that the music suddenly stopped).

```
<emotionml>
    <emotion>
        <category set="everydayEmotions" name="satisfaction"/>
        <link role="experiencedBy" uri="file:john.vcard"/>
        <link role="expressedBy" uri="file:johnsParty.avi" start="10s" end="15s"/>
        <link role="triggeredBy" uri="file:guestList.xml#numberOfGuests"/>
            <!-- many people have come to John's party -->
```

```
    </emotion>
    <emotion>
        <category set="everydayEmotions" name="surprise"/>
        <link role="experiencedBy" uri="file:john.vcard"/>
        <link role="expressedBy" uri="file:johnsParty.avi" start="10s" end="15s"/>
        <link role="triggeredBy" uri="file:johnsParty.avi" start="8s" end="10s"/>
            <!-- John realises that the music suddenly stopped -->
    </emotion>
</emotionml>
```

The following example describes an expression in which the "face" mode conveys joy and the "voice" modality irritation. In each **<emotion>** there is only a single **<link>**, which has the default role "expressedBy". The experiencing subject is not explicitly stated, so the markup is with respect to the person whose expression is contained in the linked resource at the given time.

```
<emotionml>
    <emotion>
        <link uri="file:johnsParty.avi" start="20s" end="23s"/>
        <category set="everydayEmotions" name="joy"/>
        <modality mode="face"/>
    </emotion>
    <emotion>
        <link uri="file:johnsParty.avi" start="20s" end="23s"/>
        <category set="everydayEmotions" name="irritation"/>
        <modality mode="voice"/>
    </emotion>
</emotionml>
```

These examples show that it is possible to encode simple co-occurrences using separate **<emotion>** elements. However, in some scenarios it may be desired to make it explicit that an emotion is complex, or, put the other way round, that an emotion is part of a complex emotion. This is not possible with the current specification. One option would be to add an additional enclosing element around the **<emotion>** tags involved. However, this solution is sub-optimal, as it would not easily allow to represent emotions that overlap in time only partly - emotions would have to be chopped into several segments, so that the overlapping part can be enclosed with a joint parent element. An alternative option to indicate the

fact that an emotion is part of a complex emotion would be an attribute providing a cross-link between **<emotion>** elements. Such cross-links could also describe the nature of the relation between the emotions, which could comprise "co-occurrs" (or more explicitly, "different cause" or "different modality" among others) or regulation (potentially "masks", "is-masked-by", etc.).

It will have to be made clear what would be the added value of such explicit labeling and who would benefit from this. One clear benefit is an increased ease in identifying emotions which are part of a complex emotion, for both humans and automatic analysers. Agreement already exists that individual **<emotion>** elements should retain their individual time stamps.

6.2 Regulation

Regulation is tightly linked to complex emotions that arise from several concurrent emotions. In particular it specifies how complex emotions are externalized: they can be superposed; but one emotion can mask another one; an emotion can be inhibited (which can correspond to a special case of masking an emotion by a 'neutral' state); an emotion can also be exaggerated, etc. Regulation can act on the display of facial expression of emotions. But it can also act at an earlier stage of the emotional process through the re-appraisal of the situation (Gross, 2001). Regulation may occur due to some socio-cultural rules (Ekman, 2003).

There are examples of use cases that may require the specification of the regulation process. Embodied Conversational Agents are virtual entities with human-like communicative qualities. They are often used in human-machine interaction to converse with a human user. To ensure socially-appropriate behaviours, ECAs may have to modulate their behaviours, such as not to be impulsive (Pelachaud et al, 2001), obey politeness rules (Niewiadomski & Pelachaud, 2007; Rehm & André, 2005), inhibit their emotions (Prendinger & Ishizuka, 2001). In some occasions, ECAs may have to put a smile as a polite sign or may have to conceal a negative emotion by a positive one. Emotional states of ECAs need to be regulated to ensure the socially-appropriate expression display.

Regulation has not been defined as a tag yet, as it is not part of the mandatory requirements for EmotionML. As representing regulation with scientifically grounded terms is very complex, we have preferred to leave aside this aspect for the moment, even though we are aware of the necessity to add it. It will be dealt with in a future revision of this specification.

6.3 Further requirements

A number of further optional requirements were left as future addenda in a prioritisation poll. These are of two main types: metadata and ontologies.

Metadata includes one local metadata item to attribute acted emotions with e.g. perceived naturalness, authenticity, and quality, and four types of global metadata, as follows. Firstly information about the person(s) involved. Depending on the use case, this would be the labeler(s), persons observed, persons interacted with, or even computer-driven agents. Secondly information about the the genre of the observed social and communicative environment and more generally of the situation in which an emotion is considered to happen (e.g. fiction (movies, theater), in-lab recording, induction, human-human, human-computer (real or simulated)), interactional situation (number of people, relations, link to participants). Thirdly, the purpose of classification, as the result of emotion classification is influenced by it. For example, a corpus of speech data for training an Embodied Conversational Agent might be differently labelled than the same data used for a corpus for training an automatic dialogue system for phone banking applications; or the face data of a computer user might be differently labeled for the purpose of usability evaluation or guiding a user assistance program - these differences are application or at least genre specific and independent from the underlying emotion model.Finally, global metadata on the technical environment, as the quality of emotion classification and interpretation, by either humans or machines, depend on the quality and technical parameters of sensors and media used (e.g. frame rate, resolution, colour characteristics of video sources; ;type of microphones, sensing devices for physiology or movement; data enhancement algorithms applied, etc.). The emotion markup should also be able to hold information on which way an emotion classification has been obtained, e. g. by a human observer monitoring a subject directly, or via a life stream from a camera, or a recording; or by a machine, utilising which algorithms.

Ontologies of emotion descriptions should provide two kinds of information: relationships between the concepts in a given emotion description, and mappings between different emotion representations.

The above-mentioned requirements have been left for future specification. The reason is two-fold: on the one hand, they do not belong to the core problem and can thus be shifted to existing, specialised markup languages (this may be the case in particular for global metadata). On the other hand, ontologies are arguably too complex to be handled at this stage. For more details, see EmotionML Requirements.

7 References

7.1 Technical references

EMMA

EMMA: Extensible MultiModal Annotation markup language version 1.0, Michael Johnston, et al., Editors. World Wide Web Consortium, 11 December 2007.

Emotion Incubator Group

W3C Emotion Incubator Group, M. Schröder, E. Zovato, H. Pirker, C. Peter, F. Burkhardt, Editors. Final Report of the Emotion Incubator Group at the World Wide Web Consortium, 10 July 2007.

EmotionML Requirements

Emotion Markup Language: Requirements with Priorities. F. Burkhardt and M. Schröder. W3C Incubator Group Report, 13 May 2008.

RDF

RDF/XML Syntax Specification (Revised), Dave Beckett, Editor. World Wide Web Consortium, W3C Recommendation 10 February 2004.

RelaxNG

RELAX NG Specification, James Clark and Makoto Murata, Editors. OASIS, Committee Specification, 2001.

RFC2119

Key words for use in RFCs to Indicate Requirement Levels, S. Bradner, Editor. IETF RFC 2119, March 1997.

RFC3986

Uniform Resource Identifier (URI): Generic Syntax, T. Berners-Lee et al., Editors. IETF RFC 3986, January 2005.

Schematron

Information technology - Document Schema Definition Languages (DSDL) - Part 3: Rule-based validation - Schematron. ISO/IEC 19757-3:2006.

SMIL Clock Value Syntax

Synchronized Multimedia Integration Language (SMIL) Version 3.0 Time Clock Syntax, S. Mullender et al., Editors. W3C Proposed Recommendation, 6 October 2008.

SSML

Speech Synthesis Markup Language (SSML) Version 1.0, Daniel C. Burnett, et al., Editors. World Wide Web Consortium, W3C Recommendation, 7 September 2004.

SSML 1.1

Speech Synthesis Markup Language (SSML) Version 1.1, Daniel C. Burnett, et al., Editors. World Wide Web Consortium, W3C Candidate Recommendation, 7 November 2008.

XML-NS10

Namespaces in XML 1.0, Tim Bray et al., Editors. World Wide Web Consortium, W3C Recommendation, 16 August 2006.

XML-NS11

Namespaces in XML 1.1, Tim Bray et al., Editors. World Wide Web Consortium, W3C Recommendation, 2006.

XML Schema
XML Schema Part 1: Structures Second Edition, Henry S. Thompson et al., Editors. World Wide Web Consortium, W3C Recommendation, 2004.

XProc
XProc: An XML Pipeline Language, Norman Walsh et al., Editors. World Wide Web Consortium, Working Draft, 14 August 2008.

XSLT
XSL Transformations (XSLT) Version 1.0, James Clark, Editor. W3C Recommendation. 16 November 1999.

W3C Datetime Note
Date and Time Formats, M. Wolf and C. Wicksteed. W3C Note, 1997.

7.2 Scientific references

Ekman, 2003
P. Ekman (2003). The Face Revealed. Weidenfeld & Nicolson, London.

Gebhard, 2005
Gebhard, P. (2005). ALMA - A layered model of affect. In Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-05). Utrecht.

Gross, 2001
Gross, J. J. (2001). Emotion Regulation in Adulthood: Timing Is Everything. Current Directions in Psychological Science, 10(6), 214-219.

Niewiadomski & Pelachaud, 2007
R. Niewiadomski, C. Pelachaud, Model of Facial Expressions Management for an Embodied Conversational Agent, ACII, Lisbon, September 2007.

Ortony et al., 1988
Ortony, A., Clore, G. L., & Collins, A. (1988). The Cognitive Structure of Emotion. Cambridge, UK: Cambridge University Press.

Pelachaud et al., 2001
C. Pelachaud, I. Poggi, B. De Carolis, A reflexive, not impulsive Agent, The Fifth International Conference on Autonomous Agents Montreal, May 2001.

Prendinger & Ishizuka, 2001
PRENDINGER H., ISHIZUKA M.: Let's talk! socially intelligent agents for language conversation training. In IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans (2001), vol. 31, pp. 465–471.

Rehm & André, 2005

REHM M., ANDRÉ E.: Catch me if you can: exploring lying agents in social settings. In Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (AAMAS) (2005), ACM Press, pp. 937–944.

8 Acknowledgments

The work leading to the present report has received numerous contributions, including discussion of requirements in various use cases, pointers to relevant scientific work, as well as the proposal and discussion of possible elements of the specification, from the following persons (in alphabetical order):

- Andrew Breen, Nuance Communications
- Dylan Evans, University College Cork
- Nestor Garay Vitoria, University of the Basque Country
- Alain Giboin, INRIA Sophia Antipolis
- Bill Jarrold, SRI International
- Kostas Karpouzis, Image, Video and Multimedia Systems Lab (IVML-NTUA)
- Myriam Lamolle, University of Paris VIII
- Hannes Pirker, Austrian Research Institute for Artificial Intelligence
- Jianhua Tao, Chinese Academy of Sciences
- Idoia Zearreta, University of the Basque Country

The editor wishes to thank the MMI group for their helpful comments, and Liam Quin for the suggestion of using XProc for validating EmotionML.